
8

Market Share Estimation: Data Mining for an Exceptional Case

Appendix 8.A Dummify PROMO_Code

```
PROC TRANSREG data= promo_ID DESIGN;  
model class (PROMO_Code / ZERO='xx');  
output out = PROMO_Code (drop = Intercept _NAME_ _TYPE_);  
id ID;  
run;
```

```
PROC SORT data= PROMO_Code; by ID;  
PROC SORT data= RAL_data; by ID;  
run;
```

```
data RAL_data1;  
merge RAL_data PROMO_Code;  
by ID;  
run;
```

Appendix 8.B PCA of PROMO_Code Dummy Variables

```
PROC PRINCOMP data= RAL_data1 n=16 outstat=coef
out=RAL_data1_pcs
prefix=PROMO_Code_pc std;
var
PROMO_Code1
PROMO_Code2
PROMO_Code3
PROMO_Code4
PROMO_Code5
PROMO_Code6
PROMO_Code7
PROMO_Code8
PROMO_Code9
PROMO_Code10
PROMO_Code11
PROMO_Code12
PROMO_Code13
PROMO_Code14
PROMO_Code15
PROMO_Code16
;
ods exclude cov corr SimpleStatistics;
run;
```

Appendix 8.C Logistic Regression YUM_3mos on PROMO_Code Dummy Variables

```
ods exclude ODDSRATIOS;
```

```
PROC LOGISTIC data=RAL_data1_pcs nosimple des outest=coef;  
model YUM_3mos =  
PROMO_Code_pc1-PROMO_Code_pc15;  
run;
```

Appendix 8.D Creating YUM_3mos_wo_PROMO_CodeEff

```
PROC SCORE data=RAL_data1_pcs predict type=parms score=coef out=score;  
var PROMO_Code_pc1-PROMO_Code_pc15;  
run;
```

```
data score;  
set score;  
estimate=YUM_3mos2;  
run;
```

```
data RAL_data1_wo_PromoEff;  
set score;  
prob_hat=exp(estimate)/(1+ exp(estimate) );  
YUM3mos_due_PROMO = prob_hat;  
YUM3mos_wo_PROMOeff = 1- prob_hat;  
run;
```

Appendix 8.E Normalizing a Variable to Lie Within [0, 1]

```
PROC RANK data=RAL_data1_wo_PromoEff normal=TUKEY
    out= X_RNORMAL ties=dense;
var YUM3mos_wo_PROMOeff;
ranks RX;
run;
```

```
PROC UNIVARIATE data=X_RNORMAL plot;
var RX;
```

```
PROC MEANS data = X_RNORMAL min max;
var RX; run;
```

```
* Subtract min. value of RX, divide by max. value of RX;
data X_RNORMAL;
set X_RNORMAL;
RXX =(RX+3.8138025)/3.9474853/1.9661347;
```

```
PROC MEANS data = X_RNORMAL min max mean;
var RXX; run;
```

```
* Center RXX at mean=0.20 fiddle with std values to yield 0<= RXX <=1;
PROC STANDARD data=X_RNORMAL mean=.2 std=.15
out=XRNORMALZ20;
var RXX;
run;
PROC UNIVARIATE data=X_RNORMALZ20 plot;
var RXX;
run;
```

```
title' MarketShare_est20=((RMarketShare+3.8138025)/ 3.9474853)/1.9661347 ';
data MKTShare_RNORMALZ20;
set X_RNORMALZ20;
MarketShare_est20=RXX;
run;
```